DOCUMENT RESUME

ED 256 172 FL 014 999

AUTHOR	Oskarsson, Mats
TITLE	Self-Assessment of Foreign Language Skills: A Survey of Research and Development Work.
INSTITUTION	Council for Cultural Cooperation, Strasbourg (France).
REPORT NO	ISBN-92-871-0358-5
PUB DATE	84
NOTE	53p.
PUB TYPE	Information Analyses (070) -- Reports - Research/Technical (143)
EDRS PRICE	MF01/PC03 Plus Postage.
DESCRIPTORS	*Adult Students; *Language Research; *Language Skills; *Language Tests; Literature Reviews; *Second Language Learning; *Self Evaluation (Individuals); Surveys; Teacher Education; Test Reliability; Test Validity

ABSTRACT

A review of research in self-assessment techniques for foreign language skills is based on a survey of the literature and the results of a survey of teachers, researchers, and administrators in adult education centers in many areas of the world. Two late 1970s surveys of self-assessment research and pilot experiments concerning the general validity and usefulness of learners' estimates are described, and more recent research is reviewed. A growing body of research on the use of self-assessment procedures for continuous evaluation purposes is also examined, and the survey results are reported. The pattern emerging is of consistent overall agreement of self-estimates and external criteria, but the need for practice in autonomous learning and self-directed evaluation is emphasized in some of the sources, and it is suggested that teacher training in these areas is a further prerequisite for effective development of self-evaluation techniques. A variety of self-assessment aids are described and illustrated, and it is noted that the kind of instrument seeming to yield the most accurate information consists of descriptions of concrete linguistic situations that the learner can interpret and evaluate in behavioral terms. Implications of the findings and areas needing further study are discussed. (Author/MSE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED256172

COUNCIL OF EUROPE

CONSEIL DE L'EUROPE

EDUCATION & CULTURE

Self-assessment of foreign language skills: a survey of research and development work

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Mats Oskarsson

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

FL01-1999

Strasbourg
1984

COUNCIL OF EUROPE

CONSEIL DE L'EUROPE

Self-assessment of foreign language skills: a survey of research and development work

Mats Oskarsson
University of Göteborg
Sweden

COUNCIL FOR CULTURAL CO-OPERATION

Strasbourg
1984

This work has been published in French under the title:
Auto-évaluation des aptitudes en matière de langues vivantes
ISBN 92-871-0357-7

Strasbourg, Council of Europe, Publications Section
ISBN 92-871-0358-5
© Copyright, Council of Europe, Strasbourg 1984
Printed in Belgium

The opinions expressed in this work are those of the authors and are not to be regarded as reflecting the official policy of the Council for Cultural Co-operation of the Council of Europe.

All correspondence concerning this report or the reproduction or translation of all or part of the document should be addressed to the Director of Education, Culture and Sport of the Council of Europe (F-67006 Strasbourg Cedex).

TABLE OF CONTENTS

		Page
1.	INTRODUCTION	4
2.	PREVIOUS SURVEYS OF SELF-ASSESSMENT RESEARCH	4
2.2	Heidt	4
2.3	Oskarsson	6
3.	RECENT RESEARCH	7
3.2	Raasch	7
3.3	Ferguson	9
3.4	Windeatt	9
3.5	Palmer and Bachman	12
3.6	Evers	14
3.7	LeBlanc and Painchaud	15
3.8	Lee and Low	16
3.9	Fok	16
3.10	Achara	17
4.	FOCUS ON CONTINUOUS ASSESSMENT -- RECENT RESEARCH	20
4.2	Eurocentre Bournemouth	20
4.3	von Elek	22
4.4	Rea	26
4.5	Heindler	27
5.	SHORT NOTICES	30
6.	SUMMARY AND CONCLUSIONS	31
References		33
Appendices		36

1. INTRODUCTION

The main purpose of the present study is to review reports on research and development work in the area of self-assessment of foreign language learning. The focus of attention is on adult learning, particularly at lower levels, and on procedures and results that are judged to be of general interest, especially in relation to the work on adult language learning carried out within the Council of Europe. (For a comprehensive description of this project, its background, development and principal results, see the Council of Europe Modern Languages, 1971-1981, with a résumé by the project adviser (J Trim). For more detailed information, see for instance Trim, 1978; Trim, Richterich, van Ek, and Wilkins, 1973; and van Ek, 1975).

Another aim of the study is to illustrate various types of materials that have been used for self-assessment purposes in teaching and research.

The study was commissioned by the Council of Cultural Co-operation of the Council of Europe and forms part of the very extensive programme of work undertaken within the Modern Languages project. The content of the report links up with the results of a previous investigation on self-assessment (described in Oskarsson, 1978 and 1981), which was also carried out within the framework of the Modern Languages Project. A brief summary of the results of this pilot study is given below (Section 2.3).

The information collected here was obtained mainly through the following activities: (a) a search of the literature, (b) the administration of a questionnaire requesting information about ongoing activities in the field, (c) participation in research projects dealing with theoretical and practical problems in the area of self-assessment, (d) a study visit to Eurocentre Bournemouth. The emphasis has been on the search of the relevant literature, whereby the information obtained through the administration of the questionnaire proved to be of considerable use. The activities specified under (c) and (d) each resulted in important insights which were used in the interpretation and evaluation of research and development reports.

I wish to take the opportunity here to thank all those who have contributed to the survey by supplying information about their various activities and by granting permission to re-produce certain materials for illustration purposes. Without their assistance my task would have been much more arduous and considerably less interesting.

2. PREVIOUS SURVEYS OF SELF-ASSESSMENT RESEARCH

2.1 Two pieces of research relevant to the aims of the present study will be briefly reviewed here. One of these is a project undertaken by Heidt (1979) as part of a larger inquiry into the characteristics and potentials of "autonomous learning" (defined as "learning carried out in the absence of a teacher"). The other is the pilot investigation mentioned in the Introduction above, in which a preliminary attempt was made at exploring possible forms of self-assessment in adult language learning. The former study was commissioned by UNESCO and the latter by the Council of Europe.

2.2 Heidt The Heidt study, which is the most comprehensive and detailed survey that we have been able to identify in our search of the literature, is similar to the one described in the present report in so far as it deals with the same basic problem and reviews the same type of educational

thinking and research. It is different, however, in that it is much more general in nature and attempts to analyse the whole concept of self-evaluation in learning ("from preschool up to adult learning"). Furthermore, the emphasis in the Heidt study is on general considerations "centering on purposes of assessment, justification of self-assessment and the qualifications of assessors" (page 2) and on a broad discussion of the role attributed to self-evaluation in different educational contexts (including formal schooling, adult learning, distance education, teacher instruction and its derivatives, and student-teacher contracting and project methods). The account of concrete research findings is given a less prominent position. The present survey differs from the Heidt study first in that it is restricted to the area of foreign language learning alone, and second in that it is primarily devoted to a description of the practical application of self-assessment procedures in research and development work and to a fairly close scrutiny of the results obtained.

The reason why empirical evidence is lacking to a large extent in the Heidt study is all too obvious (and is pointed out passim in the report):

"Contrary to expectations, the extensive literary search revealed only a disappointingly small number of reports and studies relevant to the purpose. Although self-evaluation, self-control, and self-assessment are frequently mentioned in recent educational literature and although it is common practice today to include self-evaluation opportunities in instructional materials, there is no equivalent research to date as to the effects of the effectiveness of such self-assessment opportunities. It has not been possible to identify a single report focusing on practical experiences in the field, and the existing relevant empirical studies are almost a negligible quantity" (page 34)

Although the survey "revealed a disappointingly small number of reports and studies relevant to the purpose", as the author puts it, the report contains a wealth of information, based as it is on a very thorough literature search. Anyone who is interested in the topic from a general point of view is therefore advised to study the report.

One of the most obvious conclusions that Heidt was able to draw on the basis of his search for practical examples of self-assessment was that "almost all books which in the widest sense of the term may be called instructional or self-instructional contain some kind of self-assessment component. In most cases each chapter, often called "unit", is followed by questions, exercises, or by a self-diagnostic test. If the reader can pass this test - and it is his own personal decision whether he does so or not - he is directed to the next chapter. Even if the self-assessment exercises sometimes take a more elaborate form, they are only a variation developed from this basic model" (page 10).

As regards language learning (pages 19-20), which is our immediate concern here, Heidt makes a reference to a Unesco investigation of foreign language materials (no source quoted) in which the following conclusions were reached:

"The majority of courses under investigation in this analysis have some kind of self-assessment and self-correction component. A second point refers to a more attitudinal problem: The materials and the contexts in which they are used expect the learner to accept responsibility for his

own learning process and at least to participate in its evaluation. The analysis stresses the difficulties in providing model answers for more complex and creative exercises. Several solutions are suggested to overcome this prob c

1. The model answer consists of a "medium" response which might be given by a native speaker.

2. The solution states the different elements and components which should be included in a correct answer. This way is also chosen by a method known as "structural communication"

3. Although for exercises of the "closed type", where only one response is possible, the learner may be given the full opportunity of self-evaluation, the open-ended type of exercise requires a kind of semi-autonomy for correction, ie a co-operation of learner and teacher.

4. Autonomy in self-testing may be enhanced by the use of computer terminals which are a distinguishing feature of all kinds of computer-managed instruction" (page 19).

In general, then, the author arrives at the conclusion that self-assessment in a wide sense (as exemplified by self-tests or self-assessment exercises with keys) is a natural characteristic of most teaching materials and courses today, and that the chief problem is our scanty knowledge of the effectiveness of self-assessment principles put into practice. Finally the author also notes the particular problems involved in using self-assessment methods for purposes of checking pronunciation. (This, of course, has been pointed out by many other authorities on language learning, for example John Trim in the preface to the study described below and by Albert Raasch, 1980 (cf section 3.2 in this report).

2.3 Oskarsson The other piece of work to be summarised here is the previous investigation on the same subject undertaken as part of the research and development programme of the Modern Languages Project (Oskarsson, 1978). The work of this pilot study was carried out and first reported to the Council of Europe in 1977. In a slightly revised form it was published by the Council of Europe in 1978 (and in 1980 by Pergamon Press). A brief résumé is given in the report on the CDCC Modern Languages Programme 1971-81 (Modern Languages, 1971-1981, pages 64-65)

The chief aim of the study was to suggest, and also to a certain extent to try out, possible forms of self-assessment for use in adult language learning. Another aim was to describe related research and current practices in the field. The work was undertaken in order to further one of the fundamental aims of the Modern Languages Project, viz to develop student-centred and motivation-based approaches to language learning in adult education contexts.

The review of the literature revealed only a limited number of reports and studies on self-assessment research, as did the later survey by Heidt (1979). Most of the practical activities in the field involved self-administration of traditional tests, mostly so called discrete-point tests focusing on vocabulary, grammar, and reading comprehension. There seemed to be few attempts at trying to develop learners' ability to assess their oral skills.

In the report suggestions were made concerning possible "instruments" to be used for the purpose of self-assessment of language skills, both with regard to overall assessment and with regard to more detailed assessment of specific skills. One type was based on specifications of objectives (in terms of topics and functions) in The Threshold Level. Others were based on descriptions of everyday language situations. A general proficiency rating scale for global self-assessment in each of the four language skills (listening, speaking, reading, writing) was also produced.

It was further suggested that more informal assessment methods should be tried out, such as assessment in simulated real-life situations, impressionistic judgements of group members on each other's performance, and assessment in authentic language situations.

Some of the ideas proposed were tested in field experiments. Briefly, the aim was to determine the degree of accuracy in adult learners' estimates of their own proficiency in a foreign language. It was hoped that the experiments would result in an indication of the general validity and usefulness of self-assessment procedures.

Groups of learners were asked to make independent assessments of their proficiency using self-rating forms of the type referred to above. Teacher judgements and formal test scores were used as control measures.

In general there was high correlation between self-assessment results and objective external criteria. The subjects were able to estimate their performance levels quite well in relation to the concrete behavioural specifications provided. It was concluded that self-assessment seemed to be a fruitful field for further research and development work. However, it was also recognised that the learner cannot be charged with the task of evaluating his learning without due preparation. It is a necessary part of the whole learning process to give the student training in self-evaluation so that he may develop his ability to make reliable and valid judgements on his performance, progress and future possibilities.

3. RECENT RESEARCH

3.1 For reasons explained above (cf Sections 2.2 and 2.3) all the research activities that came to our attention are fairly recent, the oldest report under review dating back to 1978. Several of the researchers quoted seem to be busily engaged in their work at the present moment (June 1983) and the indication is that the topic of self-assessment (variously termed "self-evaluation", "self-rating", "self-testing", "self-appraisal", "self-control" etc) has only just begun to expand as a distinct field of interest in language testing and evaluation. The Modern Languages Project of the Council of Europe, firmly dedicated as it is to developing a learner-centred, needs-oriented and motivation-based language learning system, seems to have played a significant part in this development.

3.2 Raasch We will start our survey by looking at a piece of research which is similar in nature to the one described above (ie the previous Council of Europe study on self-assessment). It is reported in Raasch (1979, 1979a, 1980) and was undertaken in the context of the Deutscher Volkshochschul-Verband (DVV) of the Federal Republic of Germany, Austria, Switzerland and Holland. The DVV is an organisation offering courses to adults, mostly at elementary levels.

The basic question that Raach has been trying to answer is this: What is the relationship between adult learners' conceptions of their own language skills and the formal test results that the same learners obtain when they sit for an exam (or certificate)? Comparisons were made of the outcome of Zertifikatsprüfungen in French and the results of answers in a questionnaire (Selbsteinschätzungsbogen) aimed at discovering the students perceived mastery of the language in specified situations. The two types of measure were directed towards the same kind of objective, which made the comparison possible ("Dieser Vergleich basiert darauf, dass sowohl die Zertifikatsprüfung wie der Selbsteinschätzungsbogen von einem gleichen zugrundeliegenden Lernziel ausgehen, das in der Zertifikatsbroschüre definiert ist und das grob lautet: Sich in den Alltagssituationen der Begegnung mit Frankreich und Franzosen sprachlich behaupten können"; Raasch, 1979a, page 2).

The following samples illustrate the type of questions that were used in the self-assessment test (cf Raasch, 1979a, pages 6-7):

8. Sie unterhalten sich bei Tisch mit einer befreundeten Familie über Alltagsfragen wie z.B Ferienpläne. Wie können Sie sich in einem solchen Gespräch sprachlich äussern?

a. fliessend, fast ohne Fehler
b. ziemlich fliessend, mit einigen Fehlern
c. stockend und mit Fehlern, aber verständlich
d. Ich kan mich in dieser Situation kaum ausdrücken.

10. Sie wollen einen Zeitungsartikel allgemeinen Inhalts der Tagespolitik lesen. Wie können Sie diesen Zeitungsartikel verstehen?

Das Verständnis macht

a. fast keinerlei Schwierigkeiten
b. gelegentlich Schwierigkeiten
c. grosse Schwierigkeiten, aber es gelingt i.a. noch
d. Das Verständnis ist mir fast unmöglich.

12. Sie wollen die Nachrichten im französischen Rundfunk hören. Wie schätzen Sie Ihre Fertigkeit ein, diese Nachrichten zu verstehen?

Das Verständnis macht

a. fast keinerlei Schwierigkeiten
b. gelegentlich Schwierigkeiten
c. grosse Schwierigkeiten, aber es gelingt i.a. noch
d. Das Verständnis ist mire fast unmöglich.

15. Wortschatz- und Grammatikfehler machen Sie im schriftlichen Ausdruck

a. in allgemeinen nicht
b. gelegentlich
c. oft, aber Sie machen sich dennoch verständlich
d. sehr oft, so dass Ihre Äusserungen oft kaum verständlich sind.

(Further examples of questions used are given in Appendix 1)

In all, 231 adult students accepted an invitation to participate in the investigation (out of a total of approximately 1,000 students approached). The questionnaire was completed in connection with the Zertifikatsprüfungen. The main results can be summarised as follows:

a. The candidates' self-assessments of their receptive skills differed from formal test results to a greater extent than did their assessments of their productive skills.

b. The candidates assessed their oral proficiency more accurately than they did their written proficiency.

c. In the field of pronunciation there was a feeling of great uncertainty among the candidates. In relation to formal test results they clearly underestimated their skills.

d. Almost half the number of candidates made accurate assessments of their own ability levels. ("Fast die Hälfte der Kandidaten hat die eigene Leistung richtig eingeschätzt..."; ibid., page 2)

On the basis of the results of this investigation Raasch recommended that self-assessment procedures should be made a regular feature of the language courses offered by the DVV, the rationale being that the learners would then be better able to see in what areas they need to concentrate their efforts. It was also argued that the practice of self-assessment makes for increased student motivation (ibid, page 4)

3.3 Ferguson Ferguson (1978) examined two tests of self-assessment of oral skills and compared them with standardised measurements of oral expression. Test 1 consisted of a list of 35 questions of the following type:

Can you order a simple meal?	Yes ☐
Can you give and take simple messages on the telepone?	Yes ☐
Can you describe the geography of a familiar location?	Yes ☐

In Test 2 the subjects (89 first year university students and 90 students at an interpreter's school) were asked to rank their speaking ability in relation to samples of student recordings representing a graded scale of language proficiency attainment.

The author found that the correlation between self-assessment scores and conventional speaking test results was .39 or lower. Good students seemed to underrate their skills, while the opposite seemed to be the case with the less proficient students.

3.4 Windeatt Providing the learner with opportunities for evaluating his own proficiency is an educational ambition which is particularly relevant in the area of self-access learning (variously termed autonomous learning, self-directed learning, self-instruction etc; for a discussion of these and similar terms, see Dickinson 1978, and Candelier et al., 1975). For such modes of learning to be really effective provision needs to be made not

only for the learning process itself, ie self-study, but also for self-evaluation of what has been studied so that the learner can monitor and adjust his efforts in accordance with his achievements. (For a discussion of the issue and for an illustration of one way of setting up and running a self-access centre, see Keeler, 1982).

Windeatt (1981) discusses the general function of self-assessment procedures within the framework of a self-access learning system. The author points out that such procedures might be expected to have the effect of making the student more fully involved in the language-learning process, "with the more specific objectives of alerting him to opportunities for monitoring his own performance and of helping him come to terms with the affective demands of language-learning by allowing him to articulate his own ideas about his needs/lacks/wants" (ibid., page 58). He goes on to say that there are a number of techniques available, including interviews and group discussions "although questionnaires seem to be particularly convenient in the context of a self-access system. These might be used, for example, at the beginning of a course to decide the overall shape of that course, or within individual units of work". (A similarly positive view of the questionnaire as a self-assessment aid is taken by Raasch, 1979): "Suivant nos expériences, le meilleur point de départ pour l'auto-évaluation est un questionnaire où sont formulés des items qui représentent, de la façcon la plus concrète possible, les objectifs d'apprentissage qu'on s'est proposés pour une partie bien déterminée du cours de langue ...", page 65).

Windeatt supplies a sample questionnaire based on a model described in Oskarsson (1978), ie the study that was initiated by the Council of Europe Modern Languages Project (cf section 2.3 above). This extract shows the types of item used:

Self-Assessment Questionnaire: Language Proficiency

---

1. How well do you know English? Put a cross on the number in the right hand column which you think most accurately describes your level in English. For example, if, under 'Listening', you decide you could understand spoken English almost as well as a native speaker, you might put a cross (X) on the 10. If you have occasional difficulties, you might put a cross on the 9. If, on the other hand, you have great difficulty in understanding spoken English, you might put a cross on a much lower number, like 2 or 3.

---

iv) READING

"I can understand *almost everything* I read in English with no difficulty"	**10**
	9
	8
	7
	6
	5
	4
	3
	2
"I can understand *almost nothing* I read in English"	**1**

---

2. Read the following descriptions of situations you might find yourself in, then put a cross on the number which you think correctly describes your performance in that situation. If you have never been in one of the situations described, try to imagine how you would manage in that situation.

---

d) Do you watch English television? YES/NO

What programmes do you watch on T.V.? What programmes have you watched this week? Did you watch any programmes last week? Did you have any difficulty understanding these programmes?

"I understood *almost everything* quite easily"	**10**
	9
	8
	7
	6
	5
	4
	3
	2
"I understood *almost nothing*"	**1**

Further examples of the materials used are given in Appendix 2.

The questionnaire was used by Windeatt with 23 university students on a course in English language and study skills. It was administered in conjunction with tests and interviews (at the beginning and end of the course) and the aim was to measure the students' own assessments of their strengths and weaknesses in the language. Briefly, the following results were obtained"

1. Only the weakest students found it difficult to use the questionnaire in spite of the fact that it was in English.

2. Individual student reactions to the procedure ranged from irritation at being asked to do "the teacher's job" to appreciation of the fact that they were given the opportunity of expressing their own opinions.

3. Judging by test results and teacher estimates most of the students seemed to overrate their ability.

4. In general the students gave themselves higher ratings at the end of the course than at the beginning.

The author also observed that the subjects tended to give themselves lower ratings for pronunciation and subject related work.

3.5 Palmer and Bachman Palmer and Bachman (1981, also reported in Bachman and Palmer, 1981a) used self-assessment (as one out of three different methods) in a so-called multitrait-multimethod investigation into the construct validity of tests of speaking and reading English as a second language. The purpose of the investigation was to find out whether speaking and reading abilities are independently measurable, i.e. whether they are in fact separate abilities (or 'traits' or 'constructs'). The self-assessment method was included because it is easily adapted to the measurement of both the reading and speaking traits. Other methods used were an oral interview method and an oral translation method.

The self-assessment tests consisted of written questionnaires in the subjects' mother tongue (Mandarin Chinese). There were two basically different types of questions. One type contained questions in which the subjects were asked what they could do with the language, for example:

> "How hard is it for you to put several English sentences together in a row?"
>
> /Impossible/Very hard/Not very hard/Very easy/
>
> "How hard is it for you to use different kinds of English with different kinds of people?"

Etc.

The other type involved questions about the subjects' perception of their general control of linguistic forms (vocabulary range, grammatical accuracy, etc), for example:

> "How many English words do you know?"
>
> /Very few/Several/A lot, but not hundred / As many as most as many as most Americans know Americans know

"How many different kinds of grammar mistakes do you make in English?"

I make grammar mistakes in almost everything	Many kinds	Only a few kinds	I almost never make grammar mistakes

Since the main aim of the investigation was to study the distinct abilities underlying skills in a foreign language we will not, in this context, go into the details of the over-all results. Briefly, the researchers obtained evidence for the distinctiveness of the reading and speaking traits, and also evidence that the translation method provided the best measure of reading skills and the interview method the best measure of speaking skills.

As regards empirical data on the self-assessment questionnaire, which is more interesting from our point of view, Bachman and Palmer found that the reliability measures (Guttman's lambda and alpha coefficients) were slightly lower for the self-rating method than for some of the other method units (subtests) and equal to others. All the measures were of acceptable reliability, however, and as high as .85-.96 for the self-rating method (Speaking and Reading sub-sections) (Bachman and Palmer, 1981a, Table 1).

Substantial intercorrelations between self-rating and the other two methods were obtained (op.cit., Table 2). Coefficients (adjusted for attenuation) ranged from .51 to .74 (all significant at $p < .01$, df=74).

In a similar study undertaken by Bachman and Palmer (1981), much the same results were obtained. This time three main components, or 'traits' were studied:

a. grammatical competence, including morphology and syntax

b. pragmatic competence, including the sub-traits vocabulary, cohesion, and organisation, and

c. sociolinguistic competence, including sub-traits such as distinguishing of registers, control of non-verbal signals, and nativeness.

Four methods were used: a writing method requiring a variety of writing tasks, a multiple-choice method, an interview method (similar to the FSI oral interview), and a self-assessment method consisting of 24 items using 4-point scales. The items required the subjects to rate their ability, difficulty, deficiency, etc. in each of the trait areas. The subjects were 116 non-native English speakers with 18 native language backgrounds.

High reliability indices (alpha coefficients) were obtained for all the measures in the study. The coefficients ranged from .86 to .92 for the self-assessment subtests (op.cit., Table 1). In other words, there was strong evidence that the self-assessment technique yielded quite consistent and homogeneous results. Furthermore, the correlations among the method factors indicated "substantial correlations between the self-rating and the other two methods" (op.cit., p. 462). Coefficients obtained ranged from .51 to .76 for all the trait areas and method sub-sections, except for self-rating of grammatical competence and multiple-choice testing of pragmatic competence which correlated .34 (op.cit., Appendix A). However, all correlations were significant at the .001 level of confidence (df = 109).

3.6 Evers In her paper "Auto-éevaluation du niveau en français langue étrangère" delivered at the AILA conference in 1981, Riet Evers described an experiment in which fairly high correlations between formal test results and self-assessment scores were obtained. The experiment (also described in Evers, 1979) was carried out at the Department of Applied Linguistics of the Catholic University of Nijmegen, Holland, and the sample consisted of 121 Dutch adults who participated in French courses. The test used consisted of sub-sections measuring oral production, listening comprehension, and reading comprehension. The self-assessment instrument was a questionnaire requesting the subjects to rate their ability on a seven-point Lickert-type scale. Two essentially different types of questions were asked. As in the Bachman and Palmer studies, one type probed the subjects' perception of their ability in general skills areas such as speaking, reading, grammar, pronunciation and orthography (global questionnaire). Example:

Comment jugez-vous votre connaissance de la grammaire française:

	très bien	bien	assez bien	moyen-nement	assez mal	très mal
- si vous parlez?	☐	☐	☐	☐	☐	☐
- si vous écoutez	☐	☐	☐	☐	☐	☐
- si vous écrivez?	☐	☐	☐	☐	☐	☐
- si vous lisez?	☐	☐	☐	☐	☐	☐

The other type of question measured the respondents' perceived mastery of the language in specified linguistic situations (situational questionnaire). Example:

Au cours d'une conversation avec un Français sur un sujet d'actualité, savez-vous lui donner votre opinion ou l'opinion de beaucoup de Néerlandais sur par ex.: les contacts sportif avec des pays tels que l'Argentine, ou le problème de l'energie nucléaire?

très bien		assez bien		assez mal		très mal
1	2	3	4	5	6	7

This self-assessment form was filled in again after two weeks in order to make it possible to determine the reliability of the self-ratings.

The table summarizes some major findings of the research:

Table

Correlations between Self-Assessment Scores and Test Results

	Language Test		
	Oral production	Listening comprehension	Reading comprehension
Self-assessment, global	.48	.45	.39
Self-assessment, situations	.65	.45	.63

The intercorrelation between the two types of measure (language test and self-ratings) was substantial although not high, and higher between self-ratings based on situational descriptions and test results than between global self-ratings and test results.

This shows that the subjects could more accurately assess their linguistic ability in relation to specifications of concrete communicative behaviour than in relation to general language skills (speaking, reading etc.) and in relation to the formal properties of language use (grammar, pronunciation, speaking etc.). Further there was a tendency for the self-ratings of listening comprehension to correlate lower with the test than did the self-ratings of speaking and reading.

Finally it was found that the subjects tended to underestimate their proficiency in Frnech. The tendency was more noticeable among the high-level subjects.

3.7 LeBlanc and Painchaud In a Canadian experiment reported by LeBlanc and Painchaud (1981) the point of departure of the research was the type of self-assessment material proposed in the previous Council of Europe study on the subject (see section 2.3 in this report). The research programme comprised two phases. In the first the main aim of the research was "to establish whether self-assessment would be a workable concept in the context of university students having to meet second language requirements to obtain their diploma" (ibid., p. 194). Five hundred students at the University of Ottawa were asked to complete a self-assessment questionnaire before taking a standardized second language proficiency test. The self-test consisted of a number of graded statements about the students' knowledge of the language, for example:

> "I understand simple instructions concerning my university life (studies, work, administration)."
>
> "I understand lectures on specialized topics close to my field of study."

The students assessed themselves on a scale ranging from Level 1 ("not at all") to Level 5 ("very well").

The correlations between proficiency test scores and self-assessment scores were reported to be "as high as .62 for listening and .68 for the total scores" (p. 194). The researchers concluded that most of the students "could properly assess their own level of knowledge and that self-assessment could be a valid and useful complement to and, under certain conditions, a possible replacement for standardized tests" (p. 194).

The second stage of the research (which we have not yet been able to document very well) involved a sample of over 1,200 students. The main aim was to investigate the relative value of two types of scale. The first type "was akin to that found in Oskarsson (1978) entitled Descriptive Rating Scale" and the other type "was constructed along the lines of Munby's ... taxonomy of language skills". The students filled in the questionnaire before taking the proficiency test. (Unfortunately we have not been able to obtain the full report of this part of the investigation and are therefore unable to quote the results.)

3.8 Lee and Low At the Language Centre of the University of Hong Kong, arts students taking an English proficiency test battery ("a reasonably direct type of language-use test") were asked to fill in self-assessment items at the end of the testing session (Lee, 1981, and Low 1981, 1982). The objectives were a. to find out how the students rated their performance in the test, b. to find out to what degree perceived performance co-varied with willingness to follow supplementary courses in English, and c. to establish the validity of the type of tests being used, particularly in relation to results in the writing sub-test.

A one-way analysis of variance with self-assessed test performance as the independent variable and actual test scores as dependent variables showed that the two measures "agreed reasonably well". The students' willingness to take supplementary courses in English was clearly related to their test performance.

Results based on the writing part of the test battery showed first that "students did feel they were able to make positive judgements about the task validity of the test" and second that "students were able to judge task validity independently of perception of their own performance on the test". The researchers concluded that student evaluation is one of several tools by which one may validate a test. "Cases where a testee feels that there is a serious discrepancy between his (or her) performance in a test and his actual ability can often be discovered by using self-rating questionnaires" (Low, 1982, p. 247).

3.9 Fok Further detailed information about the Hong Kong experiments is offered in a paper by Angela Fok (1981). In the introduction the author makes the point that self-assessment is an important aspect of evaluation in language learning, notably if we regard tests as having the function of indicating to the students the required target performance and not only as having the function of ranking students, selecting candidates and investigating whether relevant learning has taken place. In other words, testing methods ought to help the students rate their own ability (implicitly or explicitly) in terms of the criteria established, and in this perspective the use of self-rating procedures is a natural development.

Fok also comments on problems encountered when eliciting student self-rating, for example reliability and validity problems, the difficulty of knowing whether or not students are using the same (or comparable) criteria to measure their performance, and the difficulty of establishing the force of extraneous factors such as parental expectations, teachers' judgements, peer group pressure, and career aspirations. However, the author concludes, "we cannot deny that self-rating actually goes on all the time. In fact, it does not only exist, it is also the deciding factor determining student motivation. In situations where courses given are remedial, and non-credit giving in nature, the part played by self-rating is even more noticeable. Students only allow themselves to learn if they feel they need teaching and if they are not reasonably certain that they need to improve on their language abilities, they will not benefit from any remedial courses given, however suitably these courses are apparently tailored for their requirements" (ibid., p.2). (The claim that self-assessment is constantly practised by the language learner has also been made by other researchers, for example Holec (1981): "Les apprenant pratiquent toujours une forme d'auto-évaluation... Ceci n'a, bien sûr, rien d'etonnant, l'évaluation interne étant partie intégrante de l'apprentissage" (p. 20) and Heidt (1979, p.6): "Informal self-evaluation -- although hardly often consciously performed -- is a basic and constant component of all intentional human activities").

Fok supplies additional data on the Hong Kong project (cf.Lee, 1981, and Low, 1981 and 1982). The proficiency test battery which was used as the criterion correlated moderately with public examinations (in the region of .6), and reliability indices were in the region of .5 (reading and listening sub-sections; the speaking and writing parts were not analysed).

On the basis of self-ratings, the proficiency test scores were divided into three groups (adequate, average and good). An F test on group variance showed that the three groups were significantly different in all four skills. However, quite low intercorrelations (r = appr. .3) were obtained between the two types of measure (self-appraisal and test performance). The author attributes part of this result to a peculiarity in the listening test. Also, as was noted above, the reliability of the test was not fully established.

A further calculation revealed that students who scored high on the test could more accurately assess their own ability than mid-level and low-level students. Interestingly enough, it was also found that students rated themselves higher in the productive skills (speaking and writing) than in the receptive skills (listening and reading). The author attributes this to the fact that unfamiliar and relatively difficult types of tasks were used in the listening and reading sub-tests.

The research framework also comprised a check on background factors affecting the outcome of self-assessment procedures. Briefly, it was found that attitude towards English and socio-economic background did not influence self-ratings at all, or very little. Past academic record and performance in the test demonstrated higher relationship with the self-assessment variable.

3.10 Achara A Thai investigation, reported by Achara (1980), also involved arts students (and political science students) at university level, and some of the results in the Hong Kong experiments were confirmed. Achara found, for instance, that students who placed themselves high on a 7-point self-rating language proficiency scale also achieved significantly higher formal test results than students who rated themselves low. It was also found that

undergraduate students rated their ability in each of the four skills to be between levels 3 and 4, whereas graduate students rated their ability to be between 4 and 5. T-tests of the mean differences revealed that they were all significantly different at the .01 level of confidence. As in the Hong Kong experiments moderate to low intercorrelations were obtained between self-ratings and formal test results.

The language proficiency scale is described as a "Rating Scale for Communicative Competence in English" in which level 3 is "the threshold level". Level descriptions were presented to the students in Thai (see Appendix 3). The formal tests used consisted of reading and listening tests, an "English Proficiency Test used for screening postgraduates of Mahidol University", a cloze test and a Thai Language Aptitude Test. The above results correspond quite well to findings in other experiments, e.g. Fok (1981) and Bachman and Palmer (1981). Slightly more unexpected is the outcome of a comparison between self-ratings and ratings by the students' teachers, based on the same "Rating Scale for Communicative Competence in English". In contrast to the low correlations obtained for the relationship between self-ratings and competence as measured by means of formal tests, self-ratings and teacher ratings coincided to a surprisingly high degree. Out of a total number of 147 corresponding estimates (of listening, speaking, and reading skills) no less than 85 (58%) were identical (i.e. same level placements). 56 corresponding estimates represented a one-level difference and only 6 represented a difference of two levels (Table, Appendix 2, in the mimeographed version of the report). Accordingly, Pearson product moment correlations ranged high (from .65 to .87). We summarize the results in the diagram below:

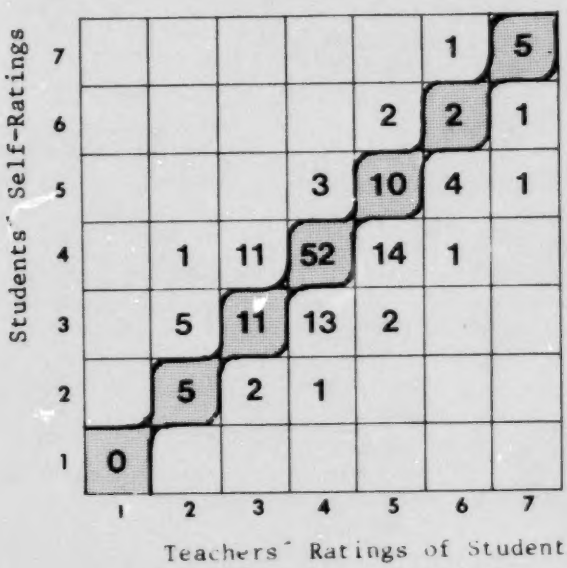


Diagram 1 The Relationship between Student Self-Ratings and Teachers' Ratings of the Students in Three Skills (The Achara Study). Total Number of Estimates: 2 x 147. Shaded Area Represents Complete Agreements in Estimates.

There was a consistent pattern of underestimation on the part of the students, with one noticeable exception: arts students tended to overrate their speaking ability. 13 students out of 22 placed themselves one "rung" too high on the proficiency "ladder" in this particular skill.

The evidence being somewhat contradictory, one may speculate about the validity of either the formal tests or the subjective assessments (student and teacher). On the one hand no statistical data on the tests are given, and the appropriateness of the tests in relation to the objectives of the research is not demonstrated or discussed. On the other hand some important information about the procedures for collecting subjective estimates is also lacking. In particular, it is never explicitly stated whether the students and teachers made their assessments completely independently of each other. One would assume that this was the case, and if so, one would probably have to give more weight to teacher estimates than to formal test scores (which, as we have seen, correlated quite low with the self-ratings).

After having read the above account of the research Dr Achara points out that the reliability of the self-assessment instrument was established in a pilot study (prior to the final experiment) by the test-retest method "yielding the reliability (stability) index of .97" and that the S's had been trained in the use of the instrument. Dr Achara also points out that the formal tests used are all standardized tests with established construct, predictive, and concurrent validity indices (and with established reliability indices).

4 FOCUS ON CONTINUOUS ASSESSMENT - RECENT RESEARCH

4.1 The research reviewed above was mainly addressed to the problem of investigating the general validity of learners' estimates. Investigations of the effectiveness of continuous "step-wise" assessment by the learners are far less common. The reason for this is, quite simply, that it is much more difficult to accommodate such activities within a research programme of ordinary and manageable proportions (unless quite substantial funds and considerable human resources are made available). However, in view of the fairly weighty evidence of the validity of learners' assessments as documented above, and in view of the body of relevant knowledge that has been accumulated in recent years, it seems that the time has now come to place the emphasis of experimental development work on the aspect of continuous evaluation (including continuous evaluation under the guidance of a teacher). Interesting efforts have been made in the past few years and we will examine some of these in this section of the report.

4.2 Eurocentre Bournemouth A good example of how the principle of continuous self-assessment may be applied in practice has been demonstrated by the Eurocentre English language school at Bournemouth. The following account of the methods used at that centre is based on descriptions in Ferris (1983), Holec (1980) and on a study visit by the author in May 1982. Information supplied by S.J. Fortescue (of Eurocentre Bournemouth) in response to our questionnaire has also been very helpful.

A language school such as Eurocentre Bournemouth constitutes a particularly suitable ground for experimentation in the field of self-assessment, mainly because students may make use of the target-language environment in order to test their language skills and in order to "calibrate" their norms of judgement when they assess their ability to cope with real-life language situations. To a certain extent these particular circumstances have been taken advantage of at the Bournemouth centre, although the emphasis has been on self-assessment in the context of a regular teaching-learning situation. The use of modern technical equipment has played an important part in the experiments.

The types of exercises employed will be briefly described here. For further details the reader is referred to the sources quoted above.

a. Laboratory follow-up to oral practice lessons. After the students have been taught a few oral practice units, they act out and record a simple transaction in pairs. The teacher then makes comments on the results and focuses on a few points relating to fluency, intonation and pronunciation. After some practice the students re-record the interaction and then compare the first recording with the second and try to assess the progress they have made. The students' assessment has been found to be "quite reliable because they can look at certain discrete points one at a time having just practised them with the teacher. They felt a sense of achievement with this activity because progress was easy to see even for the weak student who may do no more than get the words more or less right in the second version but may have been inarticulate in the first version. They did become aware of some of the criteria involved in assessing their own performance" (D. Ferris, in Holec, 1980, Appendix II).

b. Oral homework

Students are given situations to prepare in groups. Example: You are in a new town. A. Ask where a place is. B. Say where the place is. A. Thank the first person. B. Answer the thanks. The teacher tries to supply guidance and the students usually write out their dialogues first, and then practise them until they feel ready to record. They then listen to their recordings, together with the teacher if necessary, trying to spot weak points, and re-record until they are satisfied with their performance.

The conclusion that could be drawn from this type of exercise was that "the reliability of the students' judgement tends to improve and become more refined with time ... The standards of performance achieved by the students were well above the level achieved when they did paired work in class when they first did a unit. They also expressed more of their personality as they could work together in private" (ibid., p. 47).

If the atmosphere in the group is friendly and supportive, tapes of oral homework may be played to the whole class. Each student assesses the performance on a four point scale : 1 = very good, 2 = good, 3 = acceptable, and 4 = needs more work. (The students indicate their assessments by holding up numbered cards, rather like in a skating competition.) In a particular case reported "about 80% of the class came up with assessments that corresponded to that of the teacher" (ibid., p. 47).

c. Progress videos

Students are "videoed" doing very short role-plays at 3-weekly intervals. They are then given the opportunity to see the videos towards the end of the term and assess their progress. (The students find it difficult to evaluate their performance immediately after completing the role play.) - This type of video is also used in teacher training at the centre. Among other things, trainees are asked to comment on the students' self-evaluation (which was based on a form covering the areas 'fluency', 'grammar', 'pronunciation', and also 'general impression', each evaluated on a five point scale).

The general experience gained from this type of self-assessment activity has been that good students tend to under-estimate themselves, whereas poor students tend to be fairly realistic. Beginners in particular have found it difficult to assess their performance immediately after performing. "For the most part, however, I agreed with the studens' assessments", one teacher reported.

d. Progress audio cassettes

Students are recorded in the language laboratory at intervals of three weeks, responding to stimuli. As in the procedure described above they have the opportunity to hear the cassettes towards the end of term in order to assess their progress.

Other activities practised at the centre include half-term and end-of-term oral tests, which involve assessment on cue-cards, and personal interviews, which involve asking and answering questions in topic areas such as employment, hobbies, opinions, and future plans. Model tapes recorded by native speakers

are used as aids. In addition to the general experiences reported above it has been noted that students need a very simple frame-work to assess their performance and progress, and that the aims of the exercise must be explained to them very clearly.

4.3 von Elek Probably the most extensive and best researched self-assessment materials available so far in the field have been developed for learners of Swedish as a second language (von Elek 1981 and 1982). This large-scale undertaking, sponsored by the National Swedish Board of Education, comprises basic research activities as well as an extensive and quite unique development programme, leading up to the public marketing of the final products for wide-scale use by the target group. In view of the comprehensive nature of this project, it will be dealt with here in rather more detail than the experiments previously described.

The aim of the project is fairly all-embracing. All the traditionally identified language skills, plus grammar and vocabularly, as well as the whole range of ability levels from zero mastery of the language to almost native competence level, have been taken into account in an effort to provide learners (adult immigrants) with a tool for self-diagnosis at any given point of study. The resultant "test battery" consists of sub-sections covering listening, speaking, reading, writing, grammar, and vocabulary. Each sub-section is divided into ten separate difficulty levels, each of which contains 25 test items (self-assessment questions). The entire test thus comprises a total of 6 x 10 self-assessment units and 6 x 10 x 25 (i.e. 1,500) test items.

The individual items consist of questions of the following type (all in the target language)

Vocabulary	Do you know the opposite of the underlined word?
Grammar	Can you detect the mistake in the sentence and correct it?
Listening	Do you know which picture the speaker is referring to?
Reading	Do you understand the text (the following phrases ...)?
Speaking	Do you know how to answer the question?
Writing	Can you write a correct sentence using the following words?

(See Appendix 4 for more examples of the types of question used.)

The student uses a three-choice answer form (with the three options "Yes, absolutely", "I think so", and "No"), starts from a comparatively low level (level 1 if he is a beginner) and proceeds upwards through the units until he reaches a point where he masters 80 per cent of the items (and no more). The procedure is repeated in each of the six "content areas", and it results in a skills profile which may take, for example, the following form (the shaded area covers units with 80 per cent affirmative responses):

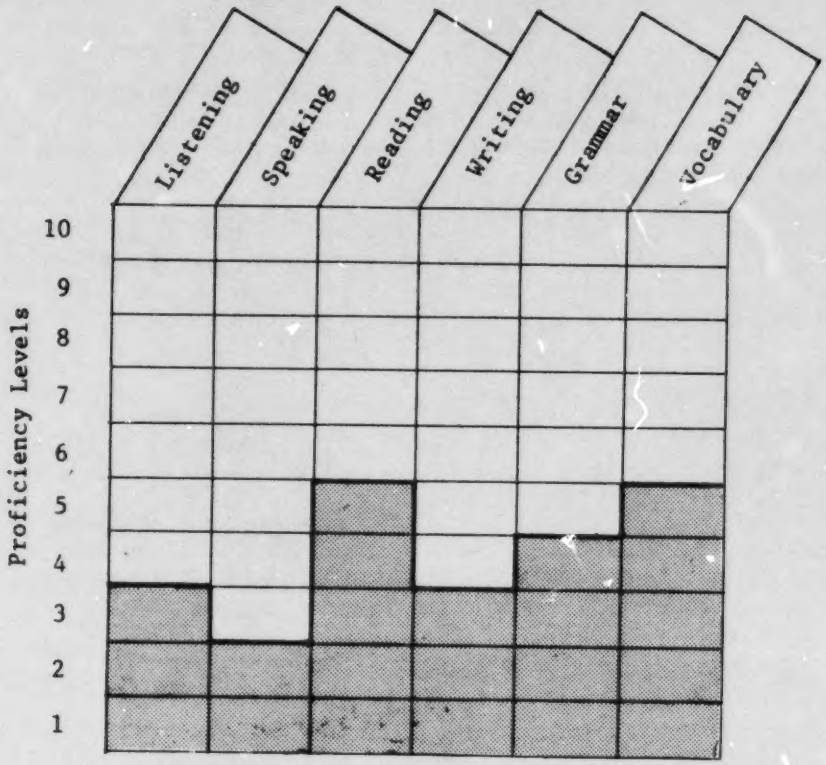


Diagram 2 An Example of a Self-Assessed Skills Profile Based on the von Elek Self-Assessment Materials

The resultant self-rated profile carries diagnostic information that the learner can use as a basis for decisions about what areas should have priority in his further work with the language. Especially in cases where there exists a target profile, a goal-line as it were, corresponding to entrance requirements and the like, such a measure of felt competence levels can give the learner considerable impetus to his efforts to achieve a particular objective.

The draft version of the test was very carefully piloted and analysed. A fairly large number of learners (adult immigrants representing various first languages, such as Finnish, Serbo-Croatian, Spanish) took the entire test battery and also participated in a follow-up check on the results, in conjunction with their teachers. The latter part of the research programme was undertaken in order to establish the reliability of the self-estimates. The student and the teacher went through all the units once more, but this time the teacher tried to probe the learner's actual skills in order to verify each statement and the teacher also did the marking (super-imposing his own marks on the answer sheet previously used by the learner). Through this procedure it was possible to ascertain the validity of student assessments in a precise and conclusive fashion. This type of validation study does not seem to have been paralleled in other experiments, at least not in

the field of foreign language learning in recent years. The reason for this is obvious. Interviewing the subjects amounted to nothing less than conducting several hundred oral tests, each involving hundreds of questions. The research task was both demanding and time-consuming, but in return it provided invaluable information on the accuracy of learners' self-estimates.

The results from the try-out are illustrated in the following scatter diagram, where we have plotted the self-assessment scores (for all sub-tests) along the vertical axis and the corresponding teacher "verdicts" on the horizontal axis (the total number of observations, i.e. corresponding student-teacher estimates, is 387):

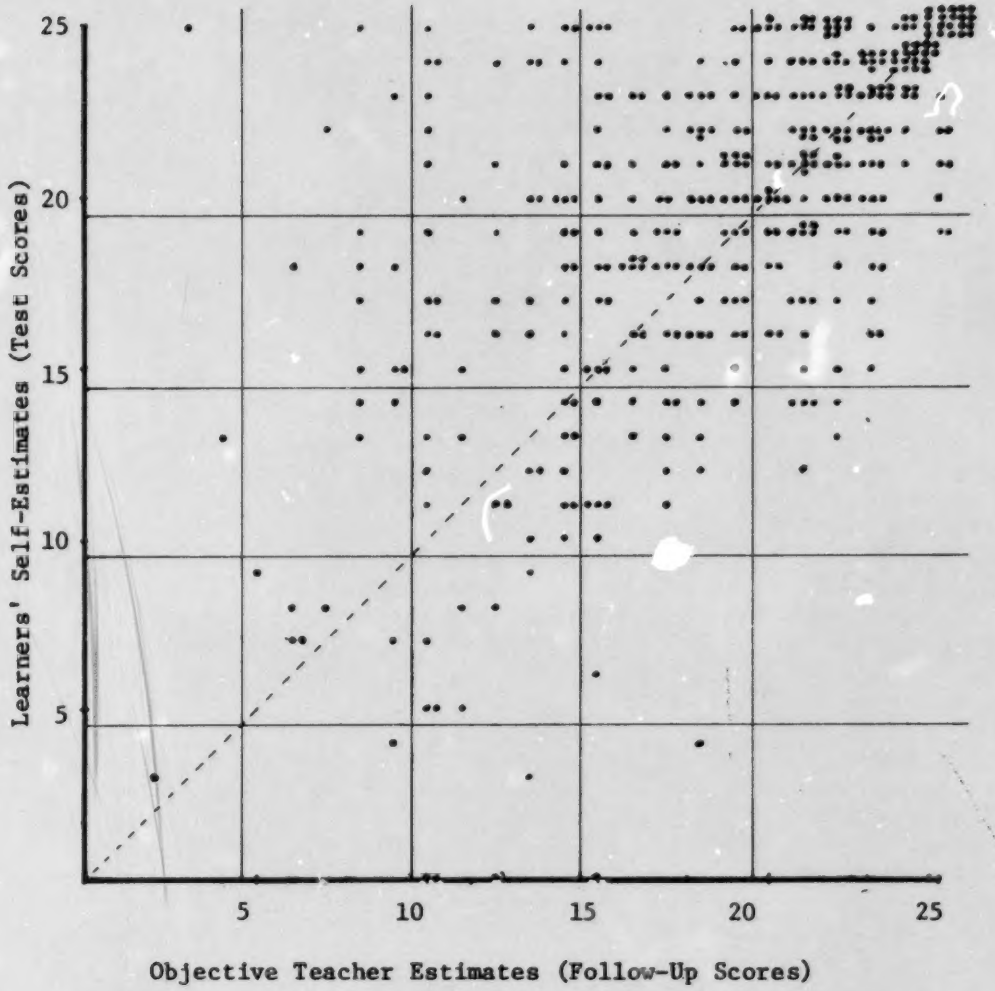


Diagram 3 A Graphic Representation of the Relationship between Self-Estimates and Objective Teacher Estimates (Follow-Up Scores) in the von Elek study

BEST COPY AVAILABLE

The correlation between the two measures is quite clear, as the diagram shows (r = .51). The correlation coefficients for the various sub-sections in the test range from .60 to .97, with the exception of the grammar parts where the coefficients were generally lower (von Elek, 1982, p. 16).

However, it is also apparent from the diagram that this preliminary version of the materials was "top-heavy" in that the scores tended to cluster at the top of the scale rather than form a nicely distributed pattern across the medium range of the continuum. The correlation coefficients must therefore be interpreted conservatively (and this is also pointed out in the report quoted above).

The number of confirmed scores, ie the number of coinciding learner and teacher estimates, was 216, which is equal to 56% of the total number of scores. As the reader will recall a similar figure for complete agreements was obtained by Achara (1981), using a 7-point assessment scale (cf Section 3.10 in this report). The number of overestimates relative to teacher verdicts, was 102 (26% of the total number of scores), and the number of underestimates was 69 (18% of the total number of scores). Thus there was a slight tendency for the subjects to overrate their language skills, contrary to what was found in the Achara and Evers studies (cf Sections 3.10 and 3.6), but in accordance with the outcome of the Windeatt study (cf Section 3.4).

Table 2 (designed for convenience of interpretation of the results) shows the number of observations in each cell of a 5 x 5 level grid. The numbers of coinciding scores are set out in the boxes forming the "perfect correlation" diagonal (shaded). The boxes to the left and to the right of the shaded diagonal represent deviations from the "true" scores (overestimates and underestimates, respectively).

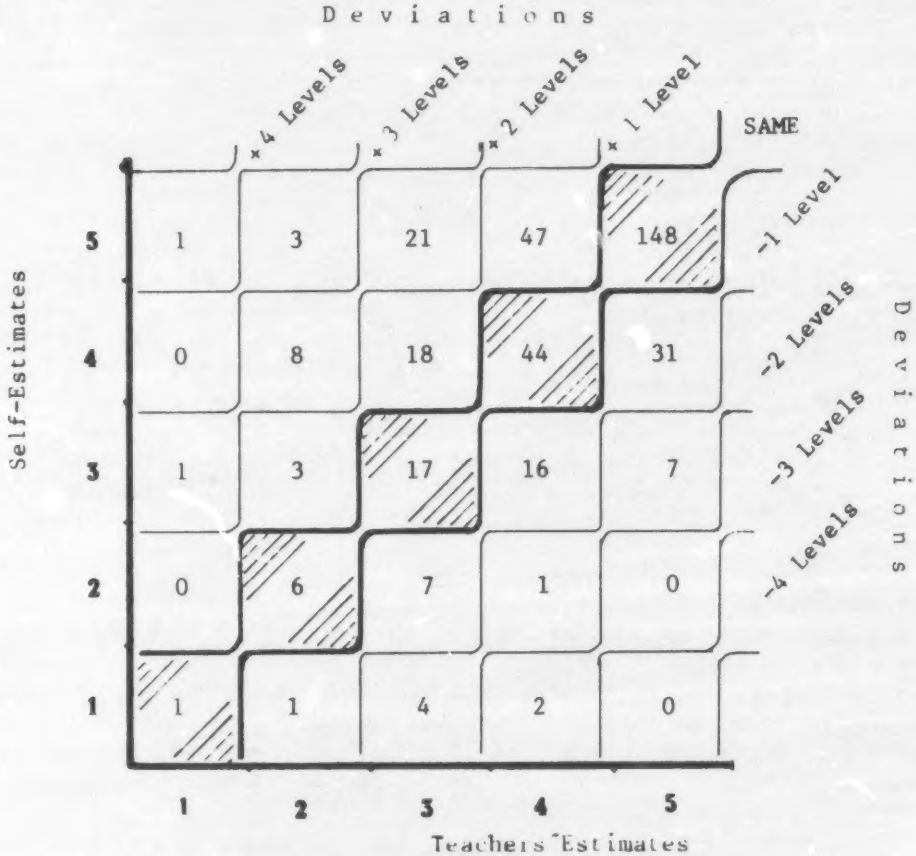


Table 2 The Relationship between Learners' Self-Estimates and Teachers' Estimates by Proficiency Scale Levels in the von Elek study.

There are relatively few cases of gross misjudgements on the part of the subjects (we may assume that the follow-up scores, i.e. teachers' estimates, are fairly dependable). As many as 381 out of 387 learner estimates (98.5%) fall within the limits of ±2 levels from the "true" scores. In other words, only 1 in 100 (or slightly more) misses the target by more than 2 levels. The great majority of subjects (88%) assess their abilities with a ±1 level error at the most.

The research not only involved examination of measurement potentials but also observation of side effects. It was found, for instance, that this exercise in self-assessment gave a boost to motivation. The following passage from von Elek (1982, p. 17) is worth quoting in full:

"One of the most important findings was that the test was very popular with the students. They generally regarded the test as a teaching aid rather than a test, and the atmosphere while they worked on it was relaxed. They were eager to find out how far they could "climb" the scale and what sort of profile their scores would yield. They were equally anxious to find out the correct answers to all items they were uncertain about or did not know. This can be seen as a beneficial spin-off effect of the test and we think it is extremely important that the curiosity and motivation this self-assessment procedure arouses should be taken advantage of in the learning process."

On the basis of the results of the first-year pilot experiments a second version of the materials was developed and field tested. The internal consistency of the battery was checked in order to ensure a balanced and progressively more difficult ordering of the units within each sub-test and also in order to ascertain comparability between the different sub-tests at each learning stage (proficiency level). The project has now (1983) been brought to a successful conclusion which, among other things, includes a decision by the National Swedish Board of Education to publish the materials on the open market (with a view to reaching as wide an audience as possible).

One of the main lessons we may learn from this project is the fact that it is quite possible to move from self-assessment in general terms, which is what most researchers in the field have been concerned with so far, to self-assessment at a more specific and detailed level -- to self-assessment of learning by units, in fact. The sample of 1,500 items constituting the test instrument that we have described can be regarded as representative of the macrostructure of language, and each of the 60 sub-tests, functioning as building blocks in this structure, can be seen as representative of a small-scale learning objective which the student himself is able to evaluate in a more or less autonomous manner. The work undertaken by von Elek seems to be very much in line with the thinking of the Council of Europe project group and can be seen as a significant contribution in a field that is regarded as a highly important feature of the Modern Languages Project -- learner autonomy and self-management.

4.4 Rea Very promising empirical data are also reported by Rea (1981). The author used a self-assessment questionnaire as a formative feature within a Communication Skills course at the University of Dar es Salaam. The general pedagogical aims of the questionnaire were to "sensitise students to language use", to encourage them to "monitor their performance", to "establish a non-competitive form of assessment", and to "increase students' awareness of their roles in the teaching and learning process". The more specific aim was to

provide the students with a yardstick by which they could measure their performance and receive immediate and continuous feedback. The questionnaire included questions about the "the estimated extent of difficulty" experienced in carrying out various language tasks within the traditional language skills areas. Answers were given on a 4-point scale ('No Difficulty', 'A Little Difficulty', 'Some Difficulty', 'A lot of Difficulty').

The instrument was administered to 200 first year students. Independently the same students were assessed by their teachers using identical assessment forms. The two sets of results were then compared with each other in a correlation study.

Although the presentation of the results of the study is not quite clear in all its details, the author was able to draw the conclusion that there was "a very high amount of agreement" between the students' assessments of their own language proficienty and the teachers' independent assessments of the same students.

Concluding the report the author states that "The self-appraisal technique does seem to provide useful information about students at the beginning of a course", and that it "provided a useful and common reference point, for teacher and student, for purposes of sensitisation and orientation, diagnosis and feedback" (p. 78). She also adds that "self-appraisal measures have inherent flexibility" because "they can be modelled on the emerging needs of the target audience and may change as and when the learner's learning environment alters".

Several examples of self-assessment materials are appended (among them some of the questionnaires presented in our previous report; see Section 2.3).

4.5 Heindler In Austria a project group under the direction of D. Heindler (1980) used ideas presented in the previous Council of Europe study on self-assessment (see Section 2.3 in this report) as a basis for evaluation (Selbsteinschätzung der Lernleistung) in the context of primary and secondary education (Gesamtschulen). The self-assessment materials formed an integrated part of a comprehensive English course ("Ticket to Britain") based on the principles expounded in the Council of Europe Threshold Level documents (but modified to a certain extent, see Heindler, 1980, pp. 36-37). The needs of lernschwache (slow learners) were given particular attention.

The following examples translated into English from the German original, show what types of item were used for the selbsteinschätzung (self-assessment): (see following page).

"You have already studied English for several weeks this term. Have you got any idea of how much you are able to say in English at this point? In the following questionnaire you can make a test of what you already know and also how well you know it. Put a cross in the appropriate box on the right hand side.

		yes	perhaps	no
2	I can introduce my friend to somebody.	☐	☐	☐
8	I can ask for somebody's telephone number.	☐	☐	☐
12	I can ask somebody to do something for me.	☐	☐	☐
20	I can tell somebody what pets my friends have got.	☐	☐	☐

If you have got more than 18 affirmative answers, we'd like to congratulate you. You have worked very well so far. If you have got 15 'yes' answers you have still done well. If you have fewer than 10 'yes' answers, you have got to figure out why this is so.

Ask your teacher to help you revise the things you have forgotten".

"How much do you know? Try to assess yourself by filling in the table.

	very well	well	to some extent	not at all
21 I understand the teacher				
25 I can speak				
28 I can work with others				

Further examples (in the original) are given in Appendix 5.

The self-tests appeared at regular intervals in the teaching materials, i.e. they were used for continuous assessment of progress during the course and not for final or global assessment. In other words, the self-tests constituted a 'formative' feature in that they were supposed to lead to modifications of the course work in accordance with the results obtained.

(Evaluation that is not used in order to provide immediate feedback on the teaching-learning process is often referred to as 'summative'.)

The teachers who participated in the development of this experimental course were asked to produce written statements of their impressions. The reports were decidedly positive. No particular effort was made to evaluate self-assessment as such (nor was this a particular objective of the research), but one teacher made the following unsolicited comment (Heindler, 1980, p. 64):

"In der Bearbeitung der Schülerselbstbeobachtungsbogen erhalten die Schüler Einblick in die Komplexität sprachlichen Geschehens. Es wurde beobachtet, dass sich Leistungsschwache Schüler in ihrem Können oft positiver und leistungsstarke Schüler skeptischer beurteilen. Hier scheint aber gerade der Angelpunkt zu sein, warum bei Schülern im Projekt kein so krasser motivationaler Unterscheid auftritt. Auch dem leistungsschwachen Schüler wird klargemacht, dass er etwas zu leisten imstande ist."

Thus, in the view of the teacher quoted, the self-tests helped the learner get a grasp of the complexity of linguistic interaction. The observation was also made that slow learners often "upgraded" themselves, i.e. over-estimated their ability, whereas the opposite was true of the more proficient students. The same effect was noted in Ferguson's Swiss study (see Section 3.3 in this report).

Continuous self-assessment as it was used by Dr Heindler and her associates is also of great interest from the point of view of organisation. Once we have been able to find a suitable format for this kind of "participatory" evaluation, we will be much closer to a situation in which the learner can and wants to assume responsibility for his own learning and thereby also for the planning and practical implementation of an effective and learner-centred language course. Thus, for instance, self-assessment instruments may easily be designed in such a way that they incorporate elements of projection into the future--on the basis of the learner's considered opinions and felt needs.

In Appendix 6 we present a draft model of a questionnaire that can be used for continuous assessment purposes. Aids of the proposed kind can be incorporated into a language course at various stages to help the learner, or learner and teacher, evaluate the results achieved and also to decide on the content of the subsequent stages of the course.

5 SHORT NOTICES

5.1 Née and Petrod (1979) administered tests in English and French to 30 business executives and secretaries and compared the scores obtained with self-assessments by the testees (on a ten-point scale). Twenty-two out of the 30 subjects placed themselves within ±1 scale level from their objectively determined proficiency level (as measured by the language tests).

5.2 A Dutch summary of the previous Council of Europe study on self-assessment (cf Section 2.3 in this report) has appeared in the Nieuwsbrief BNVU (Liemberg, 1979). Examples of experimental materials, partly translated into Dutch, were shown. The material has also been described (and in parts translated into German) by professor Fritz Satzinger in the Schriftenreihe des Verbandes Wiener Volksbildung, 3 (Fremdsprachen in der Erwachsenenbildung), pp. 57-61. A reference and a sample scale has appeared in Valette (1981).

5.3 Self-assessment in the context of "advanced conversation" courses at the University of Northern Iowa is discussed by König and Vernon (1982). The authors have found it difficult to design testing instruments and they recognise that their techniques tend to be subjective. Therefore, they point out, "In order to balance the instructor's subjectivity in the evaluation process we depend heavily on frequent self- and group-peer evaluation. The student is asked not only to give him- or herself and group members a letter grade, but also to substantiate the grades by stating a rationale. As a follow-up, the instructor meets with each individual and, using all the evaluation data, discusses current performance, any problems, and, of course, the grade" (p 21).

5.4 In Tanzania self-assessment has been adopted as one of the headings under which classroom teaching by teacher trainees is assessed (Sallie Buchanan; personal communication). Students are encouraged and trained to identify weak and strong points in their own teaching practice, and also to try to analyse their performance on the basis of questions such as: What was right/wrong with the lesson? Did the pupils learn what you were teaching? Did you hear yourself make any mistakes? Can you think of what you could do to make the lesson better? The students are also encouraged to ask themselves what they learn from demonstration lessons they attend. Self-assessment accounts for a portion of the final grades on which the student's teaching certificate is awarded.

5.5 Further examples of the use of self-rating methods in teacher training are given in Tuttle (1979), Heidt (1979), Soule-Susbielles and Weiss (1981), and Bailey (1982).

5.6 Plans for using the methods suggested in our previous report (cf Section 2.3) are reported from the Department of Foreign Languages, Tallinn Polytechnic Institute, USSR (Valmar Kokkota; personal communication). The research is planned to involve the comparison of students' self-assessment of "four parameter groups (proficiency, abilities, interest/motivation and personality) with evaluation of the same parameters by fellow students and by teachers".

6 SUMMARY AND CONCLUSIONS

The purpose of the report is to review research and development work in the area of self-assessment of foreign language learning, particularly in the adult sector. The information is mainly based on a search of the literature and on the responses to a questionnaire directed to teachers, researchers, and administrators at adult education centres.

In the first part of the report two previous surveys of self-assessment research are described. Both were published towards the end of the 1970's and showed much the same result, namely that very few empirical studies had been undertaken within this particular field of foreign language testing and evaluation. Pilot experiments concerning the general validity and usefulness of learners' estimates were described.

The second part of the report reviews recent research and development work. Most of the studies that we have been able to track down deal with the fundamental question of validity and reliability, i.e. the extend to which learners' estimates of their own language skills are consonant with independent "objective" criteria such as test results and teacher evaluation. A growing body of research on the use of self-assessment procedures for continuous evaluation purposes is also examined.

The survey shows that most of the research in the field has been carried out with students at the secondary and tertiary levels, and in the majority of cases with students of English as a foreign language. Other languages represented are French, and, in one case, Swedish (for adult migrants). Results are reported from many different parts of the world.

The general pattern of results that emerges is one of fairly consistent overall agreement between self-estimates and external criteria. Several of the studies under review contain comparisons between the two types of measure, usually in the form of calculations of Pearson product-moment correlation coefficients. Values ranging from .50 to .60 are quite common, and higher ones are not uncommon. This means, in effect, that a set of self-estimates, e.g. in the form of answers to a questionnaire based on descriptions of concrete linguistic situations, tends to carry about the same weight as a sub-test in a standardized test battery of the ordinary type. However, some researchers also point out that a given self-assessed score may be affected by errors having to do with past academic record, career aspirations, peer-group or parental expectations, lack of training in self-study and self-management etc. The need for <u>practice</u> in autonomous learning and self-directed evaluation is emphasised in some of the sources quoted. We may add that <u>teacher training</u> in these areas is a further prerequisite for the effective development of student-centred evaluation techniques.

Various types of self-assessment aids are described and also illustrated (in the Appendices). Some of them have been used as general assessment instruments, essentially for research purposes, others have been used for purposes of continuous assessment in ordinary teaching-learning situations. The kind of aid that seems to yield the most accurate information consists of descriptions of concrete linguistic situations which the learner can interpret and evaluate in behavioural terms. Higher correlations were obtained between test results and self-assessments based on such situational descriptions than between test results and <u>global</u> self-assessments (i.e. assessments in terms of "macroskills" such as 'Writing' or 'Understanding a native speaker').

In at least four of the studies examined in detail the researchers found that the more proficient students tended to underestimate their ability and skills. Reported cases of overestimation tended to involve weak students to a greater extent than good students.

Two areas of language evaluation seem to present particular problems to many learners, namely pronunciation and grammar. Low ratings as well as low correlations with objective test results are reported. On the other hand, learners appear to find it comparatively easy -- or at least easier -- to appraise their communicative skills, as evidenced by a number of the reports reviewed.

REFERENCES

Achara Wangsotorn (1981) Self-Assessment in English Skills by Undergraduate and Graduate Students in Thai Universities. In Read, J. (Ed.), Directions in Language Testing. Singapore University Press, RELC, pp 240-260. Also described under the title of "Self-Assessment of Language Proficiency", SEAMO Regional Language Centre, Sinagpore, 1980 (mimeo).

Bachman, L.F., and Palmer, A.S. (1981) A Multitrait-Multimethod Investigation into the Construct Validity of Six Tests of Speaking and Reading. In Palmer, A.S., Groot, P.J.M. and Trosper, G.A. (Eds.). The Construct Validation of Tests of Communicative Competence. Teachers of English to Speakers of Other Languages (TESOL), Washington.

Bachman, L.F., and Palmer, A.S. (1982) "The Construct Validation of Some Components of Communicative Proficiency". TESOL Quarterly, Vol. 16, No. 4, pp. 449-465.

Bailey, Kathleen, (1982) Oral Proficiency Testing of Non-native Speaking Teaching Assistants: Issues in Criterion-related Validity. Paper read at the 1982 TESOL Conference. The Monterey Institute of International Studies, Monterey, California (mimeo).

Candelier, M., Chaix, P., Coste, D., O'Neil, C., Regourd, A. (1975) A Critical Analysis of Forms of Autonomous Learning (Autodidaxy and Semi-Autonomy) in the Field of Foreign Language Learning: Report and Proposals. UNESCO, Division of Methods, Materials and Techniques, Paris.

Cox, Barbara et. al. (1974) New Approaches to Bilingual Bicultural Education: Self-Assessment Units. Dissemination Centre for Bilingual Bicultural Education, Austin, Texas.

van Ek, Jan A., (1975) The Threshold Level, Council of Europe. Published 1980 as Threshold Level English, Pergamon, Oxford.

Evers, Riet (1979) Zelfbeoordeling var taalvaardigheid frans. Department of Applied Linguistics, The Catholic University of Nijmegen, The Netherlands (mimeo). 20 pp.

Evers, Riet (1981) Auto-évaluation du niveau en français langue étrangère. Paper read at the Sixth International Congress of Applied Linguistics, Lund, Sweden (mimeo).

Ferguson, Nicolas (1978) "Self-Assessment of Listening Comprehension", IRAL, Vol. 16, No. 2, pp. 149-156.

Ferris, Dan (1983) The influence of the continuous self-evaluation of oral skills on language learning methodology. In Coste et al Contributions to a renewal of language learning and teaching. Council for Cultural Co-operation, Council of Europe, Strasbourg, pp. 171-177.

Fok, A.C.Y.Y. (1981) Reliability of Student Self-Assessment. The Language Centre, University of Hong Kong (mimeo).

Heidt, Erhard (1979) Self-Evaluation in Learning: A report on Trends, Experiences, and Research Findings. UNESCO, Division of Structures, Content, Methods, and Techniques of Education, Paris.

Heindler, Dagmar (1980) Englisch an Gesamtschulen (3. Projekbericht). Bundesministerium für Unterricht und Kunst, Zentrum für Schulversuche und Schulentwicklung, Klagenfurt, Österreich.

Holec, Henri (1980) Autonomy and Foreign Language Learning, Council of Europe. Published 1981 by Pergamon, Oxford.

Holec, Henri (1981) "Plaidoyer pour l'auto-évaluation" Français dans le Monde, 165, pp. 15-23.

Keeler, Stephen (1982) "Practising What We Teach: Teaching Teachers about Self-Directed Learning through the Integrated Use of Self-Access Environments in the Teacher Training Course" System, Vol. 10, No. 3, pp. 259-268.

Kreeft, H., and Rijlaarsdam (1980) "Zelf- evaluatie in het moedertaalondervijs", Levende Talen, No. 352, May, pp. 397-417.

König, F.H., and Vernon, N.D. (1982) Advancing "Advanced Conversation". In Garfinkel (Ed.), ESL and the Foreign Language Teacher, National Textbook Company.

LeBlanc, Raymond, and Painchaud, Gisèle (1981) Self-Assessment Scales: An Investigation into Their Validity. In Sigurd, B. and Svartvik, J. (Eds.) AILA 81: Proceedings I, University of Lund, Sweden, pp. 194-195 (abstract).

Lee Yick Pang (1981) Report on the Use of Student Self-Assessment in the Testing Programme of the Language Centre of the University of Hong Kong. The Language Centre, University of Hong Kong, (mimeo).

Liemberg, Ellie (1979) Zelf-evaluatie. Nieuwsbrief BNVU (Bond van Nederlandse Volksuniversiteiten), No. 2, Dec. 1979, pp. 7-9.

Low, Graham D. (1981) Report on the Use of Student Feedback in a Testing Programme. The Language Centre, University of Hong Kong (mimeo).

Low, Graham D. "The Direct Testing of Academic Writing in a Second Language" System, Vol. 10, No. 3, pp. 247-257.

Modern Languages (1971-1981) (1981): Report presented by CDCC Project Group 4, with a résumé by J.L.M. Trim, Project Adviser, Council of Europe, Strasbourg.

Née, C. and Pétrod, M. (1979) Sammenligning mellem resultater af selvvurdering og objective tests for 30 testede i engelsk og fransk (A Comparison of the Results of Self-Assessment and Objective Tests with 30 Candidates in English and French). Selskabet for efteruddannelse, Palægade 4, Copenhagen (mimeo), 18 pp.

Oskarsson, Mats (1978) Approaches to Self-Assessment in Foreign Language Learning. Council of Europe. Published 1980 by Pergamon, Oxford.

Oskarsson, Mats (1981) Subjective and Objective Assessment of Foreign Language Performance. In Read, J, (Ed.), Directions in Language Testing. Singapore University Press, RELC.

Oskarsson, Mats (1981) "L'auto-évaluation dans l'apprentissage des langues pour adultes: quelques résultats de recherches" Etudes de linguistique appliquée, Vol. 41 (Jan.-March), pp. 102-115.

Palmer, A.S., and Bachman, L.F. (1981) Basic Concerns in Test Validation. In Alderson, J.C., and Hughes, A. (Eds.), ELT Documents 111 - Issues in Language Testing. The British Council, London.

Raasch, Albert (1979) "S'auto-évaluer -- un néologisme?" Le Français dans le Monde, 149, nov/déc., pp. 63-67.

Raach, Albert (1980) "L'auto-évaluation dans l'enseignement des adultes" Recherches et échanges, 5, No. 1, pp. 85-99.

Raach, Albert (1979a) "Selbsteinschätzungstests für den Lerner -- eine Hilfe für den Französischlehrer" Zielsprache Französisch, 1, pp. 1-7.

Rea, Pauline M. (1981) "Formative Assessment of Student Performance: The Role of Self-Appraisal" Indian Journal of Applied Linguistics, Vol. VII, No. 1 & 2, pp. 66-88.

Soule-Susbielles, N., and Weiss, F. (1981) "Un outil d'observation: la grille linéaire", Langues Modernes, 75, 2, pp. 197-207.

Trim, John, L.M. (1978) Some possible lines of development of an overall structure for a European unit/credit scheme for foreign language learning by adults. Council of Europe. Published 1980 as Developing a Unit/Credit Scheme of Adult Language Learning. Pergamon, Oxford.

Trim, J.L.M., Richterich, R., van Ek, J.A., and Wilkins, D.A. (1973), Systems Development in Adult Language Learning. Council of Europe, published 1980 by Pergamon, Oxford.

Uren, Ormond (1975) "Thoughts on Language Testing" Audio-Visual Language Journal, Vol. 12, No. 3, pp. 135-140.

Valette, Rebecca, M. (1981) "The Evaluation of Second Language Learning". In Alatis, J., Altman, H.B., and Alatis, P. (Eds.). The Second Language Classroom: Directions for the 1980s, Oxford University Press.

von Elek, Tibor (1981) Självbedömning av Färdigheter i svenska som andra språk (Self-Assessment of Swedish as a Second Language: University of Gothenburg, The Language Teaching Research Centre, Work Paper No. 30.

von Elek, Tibor (1982) Test of Swedish as a Second Language: An Experiment in Self-Assessment. University of Gothenburg, The Language Teaching Research Centre, Work Paper No. 31.

Windeatt, Scott (1981) A Project in Self-Access Learning for English Language and Study Skills. Practical Paper in English Language Education (University of Lancaster), 3, pp. 43-82.

APPENDICES

1 Excerpt from a self-assessment questionnaire used by Raasch (1979a)

2 Samples of self-assessment materials used by Windeatt (1981)

3 Self-rating scale used by Achara (1980)

4 Sample items from a self-assessment test developed by von Elek (1982)

5 Examples of self-assessment materials used by Heindler et al (1980)

6 A proposed form for continuous self-assessment

BEST COPY AVAILABLE

APPENDIX 1

Excerpt from a self-assessment questionnaire used by Raasch (1979a). Cf Section 3.2 in this report.

Stellen Sie sich bitte vor, daß Sie in folgenden Alltagssituationen der Begegnung mit Frankreich und mit Franzosen auf Ihre französischen Sprachkenntnisse angewiesen sind. Wie schätzen Sie Ihre sprachlichen Kenntnisse und Fertigkeiten in diesen Situationen ein? Kreuzen Sie die entsprechende Antwort an.

1. Sie wollen in einem kleinen Geschäft in Frankreich einkaufen.

   Das Verkaufsgespräch führe ich
   a. fließend, fast ohne Fehler
   b. ziemlich fließend, mit einigen Fehlern
   c. stockend und mit Fehlern, aber verständlich.
   d. Ich kann mich in dieser Situation kaum ausdrücken.

2. Sie wollen in Frankreich telefonisch eine Eisenbahnauskunft einholen. Wie können Sie sich in diesem Gespräch ausdrücken?
   a. fließend, fast ohne Fehler
   b. ziemlich fließend, mit einigen Fehlern.
   c. stockend und mit Fehlern, aber verständlich.
   d. Ich kann mich in dieser Situation kaum ausdrücken.

3. In diesem Telefongespräch mit dem Auskunftsbeamten müssen Sie u. a. eine Reihe von Zahlen (Uhrzeiten) verstehen. Wie schätzen Sie Ihre Fertigkeit ein, diese Zahlen zu verstehen?

   Das Verständnis macht
   a. fast keinerlei Schwierigkeiten
   b. gelegentlich Schwierigkeiten
   c. große Schwierigkeiten, aber es gelingt i. a. noch.
   d. Das Verständnis ist mir fast unmöglich.

4. Sie wollen einen Brief an ein Hotel schreiben, um ein Zimmer für Sie und Ihre Familie reservieren zu lassen. Wie formulieren Sie diesen Brief?
   a. flüssig und fast fehlerfrei
   b. ziemlich flüssig, mit einigen Fehlern
   c. unsicher und mit Fehlern, aber verständlich
   d. Ich sehe mich zum Schreiben dieses Briefes kaum in der Lage.

5. Sie wollen Ihren französischen Bekannten Geschichten aus Ihrem Lehrbuch erzählen, die Sie im Kurs durchgenommen haben. Wie schätzen Sie Ihre Leistungen ein?

   Ich erzähle diese Geschichten
   a. fließend, fast ohne Fehler
   b. ziemlich fließend, mit einigen Fehlern
   c. stockend und mit Fehlern, aber verständlich.
   d. Ich kann mich in dieser Situation kaum ausdrücken.

6. Sie wollen sich im Gespräch mit französischen Freunden nach sozialen und politischen Problemen im Nachbarland erkundigen. Wie können Sie Ihre Fragen formulieren?
   a. fließend, fast ohne Fehler
   b. ziemlich fließend, mit einigen Fehlern
   c. stockend und mit Fehlern, aber verständlich.
   d. Ich kann mich in dieser Situation kaum ausdrücken.

APPENDIX 2

Samples of Self-Assessment Materials used by Windeatt (1981).
Cf Section 3.4 in this report.

SELF-ASSESSMENT QUESTIONNAIRE: LANGUAGE PROFICIENCY

---

1. **How well do you know English? Put a cross on the number in the right hand column which you think most accurately describes your level in English. For example, if, under 'Listening', you decide you could understand spoken English almost as well as a native-speaker, you might put a cross (X) on the 10.
If you have occasional difficulties, you might put a cross on the 9.
If, on the other hand, you have great difficulty in understanding spoken English, you might put a cross on a much lower number, like a 2 or 3.**

---

i) LISTENING

"I understand *almost everything* I hear in English"	**10**
	9
	8
	7
	6
	5
	4
	3
	2
"I understand *almost nothing* I hear in English"	**1**

---

ii) SPEAKING

a.

"I can say *almost everything* I want in English, without difficulty"	**10**
	9
	8
	7
	6
	5
	4
	3
	2
"I can say *almost nothing* in English"	**1**

b.

"Other people seem to *understand almost everything* I say in English"	**10**
	9
	8
	7
	6
	5
	4
	3
	2
"Other people seem to *understand almost nothing* I say in English"	**1**

BEST COPY AVAILABLE

BEST COPY AVAILABLE

vii) VOCABULARY

"I *rarely* come across English words which are new to me"	10
	9
	8
	7
	6
	5
	4
	3
	2
"I find most *English words difficult* to understand"	1

---

2. Read the following descriptions of situations you might find yourself in, then put a cross on the number which you think correctly describes your performance in that situation. *If you have never been in one of the situations described, try to imagine how you would manage in that situation.*

i) LISTENING

a) Have you lived in lodgings with a landlord or landlady in England? YES/NO

Think of some occasions when your landlady has talked to you. How well do you understand what she says to you? Does she have to repeat or explain things to you?

"I understand *almost everything* quite easily"	10
	9
	8
	7
	6
	5
	4
	3
	2
"I understand *almost nothing*"	1

---

b) Do you go shopping in England? YES/NO

When did you last go shopping? Did you only to into supermarkets, or did you go into a shop where you had to ask for things? Did you understand everything the shop assistant said to you, or did he or she have to repeat or explain things? Do you usually understand everything that is said to you in shops?

"I understand *almost everything* quite easily"	10
	9
	8
	7
	6
	5
	4
	3
	2
"I understand *almost nothing*"	1

BEST COPY AVAILABLE

APPENDIX 3

การพูด

Self-rating scale used by Achara, 1980 (cf Section 3.10 in this report). An excerpt from the original (in Thai) and an English translation of the complete Speaking part.

☐ ระดับที่ ๑

ไม่มีความรู้ภูไวยากรณ์ภาษาอังกฤษ สำเนียงพูดเป็นสำเนียงแบบไทยล้วน สามารถที่จะพูดบางข้อความหรือบางคำได้ในวงจำกัด เช่น พูด Good morning,Goodbye ได้ โดยที่เจ้าของภาษาผู้คุ้นเคยกับการใช้ภาษาของคนต่างชาติพอที่จะเข้าใจได้ และในการใช้ภาษาอังกฤษนั้น ๆ นี้ จำเป็นต้องใช้ท่าทางมือไม้ประกอบเพื่อช่วยให้ผู้ฟังเข้าใจ

☐ ระดับที่ ๒

มีความสามารถอ่านออกเสียงได้ พอที่จะพูดภาษาอังกฤษได้ โดยอาศัยการท่องจำเพื่อบอกความต้องการของตน โดยแทบไม่มีการเปลี่ยนแปลงคำที่ท่องจำไว้ ไม่สามารถใช้ภาษาเพื่อการสื่อความหมายหรือสร้างสรรค์ได้ มีการพูดติดตะกุกตะกัก มีที่ผิดไวยากรณ์และการออกเสียงบ่อยครั้ง ต้องพูดไปแก้ไขไป และเมื่อมีผู้มาพูดด้วยจะมีการตอบที่ช้ามาก ผู้ที่พูดด้วยต้องคอยอธิบายช่วยให้พูดต่อไปได้

Speaking

☐ Level 1

Use Thai intonation to utter certain formulaic expressions, such as "good morning" and "goodbye". To understand these utterances, the hearer has to be used to Thai intonation. Have no knowledge of English grammar.

☐ Level 2

Able to read English orally and utter memorized utterances to express one's needs. Unable to communicate freely. Commit grammatical and phonological errors. Need to repeat oneself several times. Very slow in answering any questions. Need much help from the other party in continuing the conversation.

☐ Level 3

Able to use simple sentences to talk about various topics. Language is still faulty and uncertain. Long pauses are needed before answering any questions. The other party has a difficult time grasping the meaning of the message.

☐ Level 4

Able to communicate about events. But still unable to speak freely enough to argue in discussions. Messages have fewer grammatical errors, but heavy Thai intonation. Pauses are needed because the speaker cannot think of the right vocabulary right off. Generally, the messages are understood by the hearer who may have to help during pauses.

☐ Level 5

Able to communicate in various situations. Capable of expressing one's views freely. Need preparations for complicated arguments or discussions. Rarely commit errors in grammar. The errors are stylistic most of the time. The other party, at times, may have to ask for clarification.

☐ Level 6

Able to speak on all topics and in all situations. Errors are very few and at times undetectable.

☐ Level 7

Have a native-like proficiency to communicate for all purposes. This level belongs to those who have been residents in countries where English is the mother tongue.

BEST COPY AVAILABLE

APPENDIX 4

Sample items, translated from Swedish into English, from a self-assessment test developed by von Elek (1982). Cf Section 4.3 in this report.

SELF-ASSESSMENT TEST

IN SWEDISH AS A SECOND LANGUAGE (SECOND VERSION)

SAMPLE ITEMS (TRANSLATION)

The test battery consists of 60 subtests or units falling into six content areas or skills (e g vocabulary, reading comprehension) on the one hand and into ten difficulty levels (1-10) on the other. Each unit can be identified by a symbol in which the letter stands for the skill tested and the digit for the level of difficulty. Thus V:1 identifies the lowest level of the vocabulary subtests and M:10 the highest of the oral units.[1] All units comprise 25 items arranged into different parts depending on the type of task they represent. The digit after each example below indicates level of difficulty.

VOCABULARY

Each of the ten units (V:1 - V:10) consists of three parts.

Part 1. Do you understand the underlined word?

Can you mend the car? (1)
He is known for his perseverance. (10)

Some other words tested in the same fashion are: *friendly (2), fetch (3), belong (5), reputation (7), exaggerate (8), bribe (9).*

Part 2. Do you know the OPPOSITE of the underlined word?

For children only. (3)
There is a shortage of teachers in this country. (9)

Some other words eliciting antonyms are: *long (1), always (2), wet (5), generous (7), optional (8), reproach (10).*

Part 3. Can you fill in the gap in this sentence?

What _____ is the dress? It's red. (3)
You shouldn't _____ to conclusions. (9)

Some other words elicited in this manner: *capital (2), blind (4), whose (5), oars (7), conscience (8), /take/ measures (10).*

GRAMMAR

Each of the ten units (G:1 - G:10) consists of three parts.

Part 1. Can you supply the correct form of the word in brackets?

/BUY/ What have you _____ ? (2)
/TO WATER/ Have the flowers _____? (9)

Some other forms to be supplied: *child/children (1), tooth/teeth (3), careful /-ly (6), bad/worse (7).*

---

1) The letters are initials of the Swedish terms for the subtests.

Part 2. Can you fill in the gap?

Is _____ any coffee left? (1)
I look forward to _____ from you soon. (8)

Some other words elicited in this manner: *where (3), is /in a hurry/ (4), /as if he/ was/were (6), since (7), the ... the /+ the comparative/ (9),*

Part 3. Can you see what is wrong about the sentence and correct the error?

We have played football yesterday. (3)
Are you teacher? (6)
Did you think it's me? (7)

LISTENING COMPREHENSION

Each of the ten units (H:1 - H:10) consists of three parts, two of which are pictorial. All items are tape-recorded.

Part 1. Can you identify the picture we are referring to?

The student hears:

He is wearing glasses. (1)
My sister is talking to her boy-friend. (1)

Part 2. Can you identify the person who would be likely to utter this sentence?

The student hears:

Sorry, I must have dialled the wrong number. (4)
It's a bit too hot for my taste. (4)

(At levels 8-10 the student's task in Part 2 was to alter a picture according to oral instructions.)

Part 3. Do you understand the sentence?

Would you mind spelling your name? (3)
You can make some real bargains if you are early for the sales. (5)
According to the news agency the negotiations had to be postponed until next week. (10)

BEST COPY AVAILABLE

BEST COPY AVAILABLE

READING COMPREHENSION

Each of the ten units (L:1 - L:10) consists of at least two of the following types of items:

1. Comprehension of definitions:

   Can you supply the missing word?

   In _____ the goal keeper is the only person who is allowed to touch the ball. (5)
   A _____ is a huge reptile. It moves awkwardly on shore but is extremely swift and dangerous in the water. (7)

2. Comprehension of non-contextual sentences:

   Do you understand the sentence?

   Closed on Saturdays. (1)
   There are two exceptions to this rule. (4)
   Due to our well-organised pre- and post-natal care the infant mortality rate is extremely low. (8)

3. Comprehension of coherent texts.

   The number, length and complexity of the texts varied considerably at different levels. Each text was followed by a number of questions concerning words, expressions or context.

ORAL PRODUCTION

Each of the ten units (M:1 - M:10) consists of at least three of the following tape-recorded parts.

1. Repetition of sentences.

   Can you repeat the sentence?

   See you tomorrow! (1)
   This is my daughter Maria. She is eleven. (2)
   If someone had told me ten years ago that I would eventually settle down in Sweden I would have taken it as a joke. (9)

2. Answering questions.

   Can you give an intelligible answer to the question?

   How long have you been in Sweden? (2)
   Can you give some reasons why people emigrate? (6)
   What kind of reform would you like to introduce in your country and why? (9)

3. Picture-cued tasks.

   Look at Picture A. What is the weather like here? (1)
   Look at sceneries A and B. Which of these places would you choose for your holidays and why? (5)
   Choose one of these pictures. Could you describe it so closely that a person who does not see the picture could get a good idea of what it represents? (10)

A

B

4. Coping with situations.

A Swede tells you that he is about to leave for your native country as a tourist. Could you advise him what he should see and do there? (4)
You are applying for a job. Could you tell the agent about your qualifications and previous experience? (6)
Someone claims that immigrants cause Sweden considerable economic losses. Could you defeat this view using objective arguments? (9)

5. Definitions, comparisons

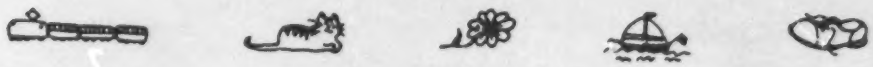
Could you explain the difference between a restaurant and a café? (4)
Could you explain the similarities and differences between a tiger and a lion? (6)
Could you explain to someone what superstition means? (10)

WRITTEN PRODUCTION

Each of the ten units (S:1 - S:10) comprises at least three of the following parts.

1. Spelling

Can you spell the following words? (2)



Can you discover the spelling mistakes in this sentence and correct them?
I am particularly interested in litterature and classical music. (10)

2. Morphology and syntax.

Can you produce correct sentences of these chains of words?
daughter / ill / cannot / go / school. (4)
grateful / information / when / product / deliver. (5)

3. Completion of sentences.

Could you finish the sentence?
Thank you for your invitation. I very much regret ... (6)
With reference to your inquiry ... (10)

4. Coping with situations.

Can you write a short autobiography? (6)
You have read an article in your native language. You think some of your Swedish friends should read it, too. Could you translate the article or write a resumé containing its essential points. (9)

5. Definitions, comparisons.

Could you explain the difference between an emigrant and a refugee in a few lines? (5)
Could you define the word hostage in a few lines? (9)

BEST COPY AVAILABLE

APPENDIX 5

Examples of Self-Assessment Materials Used by Heindler (1980) et al in the Experimental English Course "Ticket to Britain" Cf Section 4.5 in this report.

NACH DEM ZWEITEN UNTERRICHTSABSCHNITT AUSFÜLLEN

Du hast schon einige Wochen in diesem Semester Englisch gelernt. Hast du eigentlich eine Vorstellung, wieviel du auf englisch schon sagen kannst? In den folgenden Punkten kannst du selbst überprüfen, was du bereits weißt, und wie genau du es noch weißt. Kreuze das entsprechende Kästchen auf der rechten Seite an.

	ja	so halbwegs	nein
1. Ich kann jemandem meinen Namen sagen.	☐	☐	☐
2. Ich kann einen Freund von mir jemand anderem vorstellen.	☐	☐	☐
3. Ich kann jemanden nach seinen Namen fragen.	☐	☐	☐
4. Ich kann fragen, wie ein Wort auf englisch heißt.	☐	☐	☐
5. Ich weiß, was ich sagen muß, wenn ich ein Wort nicht verstanden habe.	☐	☐	☐
6. Ich kann 10 Gegenstände im Klassenzimmer benennen.	☐	☐	☐
7. Ich kann bis 30 zählen.	☐	☐	☐
8. Ich kann jemanden nach seiner Telefonnummer fragen.	☐	☐	☐
9. Ich kann angeben, welche Farbe bestimmte Gegenstände haben.	☐	☐	☐
10. Ich kann sagen, wo sich Dinge befinden.	☐	☐	☐
11. Ich kann fragen, wem etwas Bestimmtes (z.B. ein Buch) gehört.	☐	☐	☐
12. Ich kann jemanden ersuchen/bitten, etwas für mich zu tun.	☐	☐	☐
13. Ich kann vorschlagen, etwas Bestimmtes zu tun. (z.B. ins Kino zu gehen)	☐	☐	☐
14. Ich kann dem Vorschlag eines anderen zustimmen.	☐	☐	☐
15. Ich kann einen Vorschlag ablehnen und auch sagen warum.	☐	☐	☐
16. Ich kann jemanden fragen, was er/sie gern tun möchte.	☐	☐	☐
17. Ich kann sagen, was ich tun werde/ zu tun beabsichtige.	☐	☐	☐
18. Ich kann jemanden fragen, ob er pets (Haustiere) mag?	☐	☐	☐
19. Ich kann jemanden auch fragen, ob er ein Haustier hat.	☐	☐	☐
20. Ich kann angeben, welche Haustiere meine Freunde haben.	☐	☐	☐

Wenn du mehr als 18 Ja-Antworten erreicht hast, müssen wir dir gratulieren. Du hast bis jetzt sehr gut gearbeitet. Auch bei 15 Ja-Antworten hast du gut gelernt. Wenn du weniger als 10 Ja-Antworten erreicht hast, müßtest du einmal genau überlegen, wieso das so ist.

Bitte Deinen Lehrer, jene Dinge zu wiederholen, die du nicht mehr kannst.

BEST COPY AVAILABLE

Example of Self-Assessment Material used by Heindler (1980) et al

NACH DEM VIERTEN UNTERRICHTSABSCHNITT AUSFÜLLEN

★ Du hast in diesem Schuljahr schon zweimal einen Selbstbeobachtungsbogen ausgefüllt, mit dessen Hilfe du deine Leistungen im Fach Englisch selbst überprüfen konntest. Wir legen dir noch ein drittes Mal so einen Fragebogen vor. Allerdings ist er noch ein wenig umfangreicher als die beiden anderen. Wieder geht es um dich und den Englischunterricht. Wir glauben, daß es für dein Lernen wichtig ist, wenn du selbst nachdenkst,

---

wo du Spaß hast,

was du nicht magst,

wo du gut bist und wo weniger gut,

was du nicht verstehst und

wo du Schwierigkeiten hast.

Was macht dir im Englischunterricht Spaß, was nicht? Überlege dir das und trage es in die Tabelle ein.

	macht mir ganz großen Spaß	mache ich eigentlich gerne	mag ich nicht
1. Mit anderen englische Szenen spielen			
2. Neue Wörter lernen			
3. Tonband (Sprachplatten) hören			
4. Sprecher üben			
5. Schreiben üben			
6. Richtige Sätze bauen lernen			
7. Mit dem Nachbarn arbeiten			
8. In einer Gruppe arbeiten			
9. Hausübungen schreiben			
10. Tests schreiben			
11.			

BEST COPY AVAILABLE

Example of Self-Assessment Material Used by Heindler (1980) et al

Was kannst du im Englischunterricht? Versuche dich selbst einzuschätzen und trage es in der Tabelle ein.	sehr gut	gut	halbwegs	nicht
20. Ich verstehe die Tonbänder				
21. Ich verstehe den Lehrer/ die Lehrerin				
22. Still lesen kann ich				
23. Vorlesen kann ich				
24. Neue Wörter merke ich mir				
25. Sprechen kann ich				
26. Rechtschreiben kann ich				
27. Die richtige Aussprache englischer Wörter kann ich				
28. Zusammenarbeiten mit anderen Kindern kann ich				
29. Spielen englischer Szenen kann ich				
30. Selbst englische Texte schreiben kann ich				
31.				

Welche von den folgenden Dingen kannst du auf englisch sagen?	Ja	So ungefähr	Nein
32. Ich kann z.B. einen Freund fragen, ob ihm ein gefundenes Buch gehört.	☐	☐	☐
33. Ich kann jemanden fragen, was er/sie gerade sucht.	☐	☐	☐
34. Ich kann herumfragen, wem ein gefundener Gegenstand gehört.	☐	☐	☐
35. Ich kann Vermutungen äußern, wem ein bestimmter Gegenstand gehört.	☐	☐	☐

BEST COPY AVAILABLE

APPENDIX 6

A proposed form for CONTINUOUS SELF-ASSESSMENT

This is a draft form that can be used as a basis for continuous assessment of the progress of learning in a language course. If used as a model the questions ought to be translated into the learners' native language (unless the learners are at an advanced stage). The "test items" ought to be sampled, rephrased, rearranged and supplemented in such a way that they give the learners an opportunity to reflect upon all the various aspects of the course in each individual case.

---

1. In the last few lessons (days, weeks) we/I have studied/ practised/worked on ...

   (a) .....
   (b) .....
   (c) .....
   (d) .....
   (e) .....
   (f) .....

   Fill in the empty spaces with topics and areas of study that are relevant in your case, for example:

   (a) pronunciation of words containing the sound /θ/
   (b) how to greet people
   (c) questions with do/does

Leave out vocabulary ("new words") here, since this is more conveniently dealt with under item 4, below.

2. How well do you master the above topics according to your own estimate?

	not at all	to some extent	fairly well	very well	completely
(a)	☐	☐	☐	☐	☐
(b)	☐	☐	☐	☐	☐
(c)	☐	☐	☐	☐	☐
(d)	☐	☐	☐	☐	☐
(e)	☐	☐	☐	☐	☐
(f)	☐	☐	☐	☐	☐

3. On reflection, to what extent do you find the above topics important in relation to your own personal needs?

	not at all important	not very important	quite important	very important	extremely important
(a)	☐	☐	☐	☐	☐
(b)	☐	☐	☐	☐	☐
(c)	☐	☐	☐	☐	☐
(d)	☐	☐	☐	☐	☐
(e)	☐	☐	☐	☐	☐
(f)	☐	☐	☐	☐	☐

4. We/I have also studied new words of the following type, or within the following subject area(s):

(a) .....
(b) .....
(c) .....
(d) .....

(Put down your native language equivalents if you find it easier.)

5. How well do you know the above vocabulary (areas) according to your own estimate?

	not at all	to some extent	fairly well	very well	completely
(a)	☐	☐	☐	☐	☐
(b)	☐	☐	☐	☐	☐
(c)	☐	☐	☐	☐	☐
(d)	☐	☐	☐	☐	☐

6. On reflection, to what extent do you find the above type of vocabulary (areas) important in relation to your own personal needs?

	not at all important	not very important	quite important	very important	extremely important
(a)	☐	☐	☐	☐	☐
(b)	☐	☐	☐	☐	☐
(c)	☐	☐	☐	☐	☐
(d)	☐	☐	☐	☐	☐

7. Summarizing the last few lessons (days, weeks) I feel I have learnt ...

nothing at all	very little	a little	a good deal	a lot
☐	☐	☐	☐	☐

8. Looking back, I realise that I ought to change my study habits/ learning style/priorities in the following way:

.....

.....

.....

.....

.....

9. I judge my weak points to be the following:

.....

.....

.....

.....

10. I would want to see instruction in the next few lessons (days, weeks) focused on the following teaching points/skills areas:

.....

.....

.....

.....

Follow-up

Discuss your assessment and your points of view with a fellow student or in a small group -- or with your teacher. Try to find out whether others think you tend to overestimate or underestimate your ability and acquired skills and then decide whether you ought to reconsider and readjust your "yardstick". Compare your subjective impressions with other criteria such as test scores, your teacher's evaluation, estimates by your fellow students.

SALES AGENTS FOR PUBLICATIONS

OF THE

COUNCIL OF EUROPE

AUSTRIA
Gerold & Co.
Graben 31
A - 1011 VIENNA 1

BELGIUM
La Librairie européenne S.A.
244, rue de la Loi
B - 1040 BRUSSELS

DENMARK
Munksgaard Export and
Subscription Service
35, Nørre Søgade
DK - 1370 COPENHAGEN K

FEDERAL REPUBLIC
OF GERMANY
Verlag Dr. Hans Heger
Herderstr. 56
Postfach 20 08 21
D - 5300 BONN 2

GREECE
Librairie Kauffmann
28, rue Stadiou
GR - ATHENS 132

ICELAND
Snaebjörn Jonsson & Co. A.F.
The English Bookshop
Hafnarstroeti 9
IS - REYKJAVIK 101

IRELAND
Stationery Office
St Martins House
Waterloo Road
IRL - DUBLIN 4

ITALY
Libreria Commissionaria Sansoni
Via Lamarmora 45
Casella Postale 552
I - FLORENCE 50121

NEW ZEALAND
Government Printing Office
Mulgrave Street
(Private Bag)
NZ - WELLINGTON

PORTUGAL
Livraria Portugal
Rua do Carmo, 70
P - 1200 LISBON

SPAIN
Mundi-Prensa Libros S.A.
Castelló 37
E - MADRID 1

SWEDEN
Aktiebolaget C.E. Fritzes
Regeringsgatan 12
Box 163 56
S - 10327 STOCKHOLM

SWITZERLAND
Buchhandlung Heinimann & Co.
Kirchgasse 17
CH - 8001 ZÜRICH

Librairie Payot
6, rue Grenus
CH - 1211 GENEVA 11

TURKEY
Librairie Hayet Kitapevi A.S.
469, Istiklâl Caddesi
Beyoglu
TR - ISTANBUL

UNITED KINGDOM
H.M. Stationery Office
Agency Section
Publications Centre
51 Nine Elms Lane
GB - LONDON SW8 5DR

UNITED STATES and CANADA
Manhattan Publishing Company
80 Brook St., P.O. Box 650
CROTON, N.Y. 10520

STRASBOURG
Librairie Berger-Levrault
23, Place Broglie
F - 67081 STRASBOURG Cedex